

## BULK FLOW OF HALOS IN $\Lambda$ CDM SIMULATION

MING LI<sup>1</sup>, JUN PAN<sup>2,1</sup>, LIANG GAO<sup>2</sup>, YIPENG JING<sup>3</sup>, XIAOHU YANG<sup>3</sup>, XUEBIN CHI<sup>4</sup>, LONGLONG FENG<sup>1</sup>, XI KANG<sup>1,5</sup>, WEIPENG LIN<sup>3</sup>, GUIHUA SHANG<sup>4</sup>, LONG WANG<sup>4</sup>, DONGHAI ZHAO<sup>3</sup>, PENGJIE ZHANG<sup>3</sup>

*Draft version July 24, 2012*

### ABSTRACT

Analysis of the Pangu N-body simulation validates that bulk flow of halos follows Maxwellian distribution of which variance is consistent with prediction of linear perturbation theory of structure formation. We propose that consistency between observed bulk velocity and theories shall be examined at the effective scale as radius of spherical top-hat window function yielding the same smoothed velocity variance in linear theory as the sample window does. Then we compared some recently estimated bulk flows from observational samples with prediction of the  $\Lambda$ CDM model we used, some results deviate the expectation at level of  $\sim 3\sigma$  but the tension is not as severe as previously claimed. We disclose that bulk flow is weakly correlated with dipole of internal mass distribution, alignment angle between mass dipole and bulk flow has broad distribution but is peaked at  $\sim 30-50^\circ$ , meanwhile bulk flow shows little dependence on mass of halos used for estimation. In the simulation of box size  $1h^{-1}\text{Gpc}$ , for a cell of radius  $100^{-1}\text{Mpc}$  the maximal bulk velocity is  $> 500\text{kms}^{-1}$ , dipoles of environmental mass outside the cell are not tightly aligned with the bulk flow, instead are located randomly around it with separation angles  $\sim 20-40^\circ$ . In the cell showing largest bulk velocity there are slightly smaller number of low mass halos, however halos inside are clustered more strongly at scales  $\gtrsim 20h^{-1}\text{Mpc}$ , which might be a significant feature since the correlation between bulk flow and halo clustering actually grows into notable beyond such scales.

*Subject headings:* galaxies:halos — large-scale structure of universe — methods: statistical

### 1. INTRODUCTION

Bulk flow refers to the apparent coherent peculiar motion of galaxies and galaxy clusters in a considerably large volume around us. In practice there are several ways to estimate bulk flows from various observation resources, such as galaxy catalogues from peculiar velocity survey (e.g. Feldman et al. 2010), compiled Type Ia supernovae data (e.g. Dai et al. 2011), galaxy clusters in combination with cosmic microwave background (CMB) observation (e.g. Kashlinsky et al. 2010). In recent there emerge also interesting new methods based on galaxy two-point correlation function (Song et al. 2011) and galaxy light (Nusser et al. 2011; Abate & Feldman 2012).

Analysis of the spiral galaxy catalogue of the SFI++ survey (Springob et al. 2007) shows that within a top-hat spherical window of radius  $40h^{-1}\text{Mpc}$  bulk flow is  $338 \pm 38\text{kms}^{-1}$  towards Galactic  $(l, b) = (276^\circ, 14^\circ)$  with  $3^\circ$   $1\sigma$  uncertainty, and then  $257 \pm 44\text{kms}^{-1}$  towards  $(l, b) = (279^\circ, 10^\circ)$  with  $6^\circ$  error within window of radius  $100h^{-1}\text{Mpc}$  (Nusser & Davis 2011). The measurement is in agreement with analysis of Sandage et al. (2010) to data consisting of supernovae, selected nearby galaxies and galaxy clusters. Feldman et al. (2010) construct a composite catalogue of galaxies with peculiar velocities measured in different surveys, including the SFI++, estimate that bulk flow within a Gaussian window of

$50h^{-1}\text{Mpc}$  is  $416 \pm 78\text{kms}^{-1}$  in direction of  $(l, b) = (282 \pm 11^\circ, 6 \pm 6^\circ)$  (also see Watkins et al. 2009).

Employing peculiar velocities of supernovae is another viable route to detect bulk flow, though such samples usually are very sparse and prone to Malmquist bias. Dai et al. (2011) fitted a bulk flow to the Union2 supernovae catalogue (Amanullah et al. 2010) of  $188^{+119}_{-103}\text{kms}^{-1}$  in the direction of  $(l, b) = (290^{+39^\circ}_{-31^\circ}, 20^{+32^\circ}_{-32^\circ})$  within redshift  $z < 0.05$ , while no significant bulk flow is detected from data at  $z > 0.05$ ; with the same data Colin et al. (2011) obtain similar estimation but with a higher median amplitude  $250^{+190}_{-160}\text{kms}^{-1}$ . However, using a different supernovae data set within redshift shell  $z = (0.0043, 0.028)$ , Weyant et al. (2011) estimates that the local flow is  $538 \pm 86\text{kms}^{-1}$  pointing to  $(l, b) = (258 \pm 10^\circ, 36 \pm 11^\circ)$ , or  $446 \pm 101\text{kms}^{-1}$  towards  $(l, b) = (273 \pm 11^\circ, 46 \pm 8^\circ)$  if a different technique employed, which is in agreement with the dipole of CMB  $(l, b) = (263.99 \pm 0.14^\circ, 48.26 \pm 0.03^\circ)$  (Jarosik et al. 2011). Jha et al. (2007) and Haugbølle et al. (2007) have the same figures with analogous uncertainty.

Availability of modern galaxy peculiar velocity data is limited to our local Universe, bulk flow at higher redshift, sometime dubbed *dark flow*, is mainly explored through kinetic Sunyaev-Zel'dovich (kSZ) effect of galaxy clusters (Sunyaev & Zeldovich 1980). Kashlinsky et al. (2011) compute kSZ signals of 771 x-ray clusters in the 7 yr WMAP CMB map, conclude that the flow at  $z \leq 0.16$  is directing to  $(l, b) = (278 \pm 18^\circ, 2.5 \pm 15^\circ)$  and then  $(l, b) = (283 \pm 19^\circ, 20 \pm 15^\circ)$  if  $z \leq 0.25$ . They further argue that the flow at these depths shall reach magnitude of  $\sim 1000\text{kms}^{-1}$  according to an earlier investigation in Kashlinsky et al. (2010). But a conflicting assertion is derived by Osborne et al. (2011) from the same CMB map in conjunction with 736 ROSAT observed clusters, which tells that there is no significant detection of kSZ effects at low multipoles and basically deny the existence of bulk flow, meanwhile in some cases the thermal Sunyaev-

jpan@bao.ac.cn

<sup>1</sup> Purple Mountain Observatory, 2 West Beijing Rd., Nanjing 210008, P. R. China

<sup>2</sup> National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Rd., Chaoyang District, Beijing 100012, P. R. China

<sup>3</sup> Key Laboratory for Research in Galaxies and Cosmology, Shanghai Astronomical Observatory, 80 Nandan Rd., Shanghai 200030, P. R. China

<sup>4</sup> Supercomputing center, Computer Network Information Center, Chinese Academy of Sciences, 4 Zhongguancun Nansijie, Haidian District, Beijing 100190, China

<sup>5</sup> Partner Group of MPI for Astronomy, PMO, 2 West Beijing Road, Nanjing 210008, China

Zel'dovich effect might induce a dipole that could be easily misunderstood as bulk flow of  $\sim 2000 - 4000 \text{ km s}^{-1}$ .

Bulk flow is a topic of long term interests to observational cosmology (see Strauss & Willick 1995, for a review of early works), to which there are special surveys dedicated (e.g. Courtois et al. 2011). However, as we can see, no consensus on amplitude, direction, convergence depth of bulk flows has been achieved currently to reconcile different measurements. Nonetheless, some authors have argued that the amplitude of their measured bulk flow is too strong over such large scales, presenting a challenge to the standard  $\Lambda$ CDM model or at least the one of 5yr WMAP parameters (Watkins et al. 2009; Kashlinsky et al. 2010; Feldman et al. 2010; Macaulay et al. 2011). Not strangely there have been cosmological models of different flavors constructed to explain such anomaly (e.g. Mersini-Houghton & Holman 2009; Afshordi et al. 2009; Wyman & Khoury 2010).

Expectation about bulk flow in  $\Lambda$ CDM universe is generally calculated with linear perturbation theory of large scale structure in that at large scales its accuracy is believed sufficient. Often the 1-D velocity variance of dark matter are quoted to compare with measurements, however we need to address here that it is the rms velocity that should be used instead of the 1-D rms velocity. Linear theory predicts that the rms bulk velocity is typically  $\sim 300 \text{ km s}^{-1}$  with  $1\sigma$  cosmic variance approximately  $200 \text{ km s}^{-1}$  for a top-hat window of radius  $60h^{-1} \text{ Mpc}$  (Mak et al. 2011). A concern is that non-linearity might not be negligible even at very large scales (e.g. Scoccimarro 2004), which could act as systematical bias to the conclusion about the consistency between model and data.

Although linear theory can predict the possibility of observing a bulk flow of particular amplitude at certain scale, several key problems yet can not be easily tackled analytically, e.g. internal properties of the volume demonstrating large bulk blow. Practically halo catalogues from N-body simulation in large box with sufficient mass resolution are best suited for such task. The reason of focusing on halos instead of dark matter is that observational objects used to determine bulk flow are galaxies and galaxy clusters which are residing in halos, and in practice the strongly non-linearity of peculiar velocities of galaxies is largely filtered out so that what contribute to bulk flow estimation is mainly the motion following their host halos (e.g. Watkins et al. 2009). In fact Bahcall et al. (1994) and Moscardini et al. (1996) have performed analysis of mock halo catalogues and obtained useful results, but their simulations are either of very low mass resolution or based on compromised simulation method. In this paper we will demonstrate our analysis of the velocity field of halos resolved from a dark matter only  $\Lambda$ CDM simulation in  $1h^{-1} \text{ Gpc}$  box with  $3072^3$  particles. The large volume and high mass resolution of our simulation enables investigating halo behaviors in detail over broad dynamic ranges superseding previous works.

In section 2 definition and basic theoretical prediction of bulk flow is introduced, which is followed by section 3 presenting measurements of bulk flow of randomly placed cells in simulation. Section 4 is devoted to analysis of special regions showing extraordinarily large bulk flow velocity. Summary and discussion is in the last section.

## 2. BULK FLOW: DEFINITION AND LINEAR THEORY

Placing a window of characteristic scale  $R$  randomly in the sample space, if  $N$  objects (galaxies or halos, in this work just

the latter) are enclosed, bulk flow of the particular volume indicated by the particular object is

$$\mathbf{V} = \sum_{i=1}^N w_i \mathbf{v}_i / \sum w_i, \quad (1)$$

in which  $\mathbf{v}_i$  is the peculiar velocity of the  $i$ th object and  $w_i$  is the weight assigned. Practically the weights could be originated from radial selection function, angular selection function, luminosity, mass and etc. In simulation we do not have those observational systematical biases, we will discuss just two weighting schemes: window function and halo mass.

Weighting by window function only is pure geometrical, which can be easily modeled. In continuous limit,

$$\mathbf{V} = \mathbf{v} \otimes W(R) = \frac{1}{(2\pi)^3} \int \mathbf{v}(\mathbf{k}) \tilde{W}(kR) d^3k, \quad (2)$$

where  $\tilde{W}(kR)$  is the Fourier transform of the window function  $W(R)$ . The mostly common are top-hat window  $\tilde{W}_{th} = 3(\sin kR - kR \cos kR)/(kR)^3$  and Gaussian window  $\tilde{W}_G = \exp(-k^2 R^2/2)$ . There are not much differences applying top-hat or Gaussian window in principle for bulk flow measurement if effective volume and depth has been taken care of. Sometimes bulk flow estimation is provided by objects within a spherical shell defined by two radius  $R_1 > R_0$ , it is easy to see that Eq. 2 applies with window function  $\tilde{W}_S = (R_1^3 \tilde{W}_1 - R_0^3 \tilde{W}_0)/(R_1^3 - R_0^3)$ .

Probability distribution function (PDF) of  $\mathbf{V}$  could be expressed as

$$p(\mathbf{V}) d\mathbf{V} = p(V) dV p(\hat{n}_V) d\hat{n}_V \quad (3)$$

where  $\hat{n}_V$  denotes the unit vector in direction of the bulk velocity. Isotropic assumption leads to  $p(\hat{n}_V) = 1/4\pi$  and  $\langle \hat{n}_V \rangle = 0$ , which ensures that  $\langle \mathbf{V} \rangle = \int V p(V) dV \int \hat{n}_V p(\hat{n}_V) d\hat{n}_V = \langle V \rangle \cdot \langle \hat{n}_V \rangle = 0$ . Integration over the angular part of Eq. 3 yields the PDF of the amplitude of bulk flow

$$p(V) dV = \left[ \int p(\mathbf{V}) d\hat{n}_V \right] V^2 dV = 4\pi p(\mathbf{V}) V^2 dV. \quad (4)$$

$\mathbf{V}$  by definition is the velocity field smoothed by the window function, once the smoothing scale is sufficiently large the distribution of  $\mathbf{V}$  shall be very close to Gaussian so that Maxwellian distribution could be invoked to model  $p(V)$  (Bahcall et al. 1994)

$$p(V) dV = \sqrt{\frac{2}{\pi}} \left( \frac{3}{\sigma_V^2} \right)^{3/2} V^2 \exp\left(-\frac{3V^2}{2\sigma_V^2}\right) dV, \quad (5)$$

where the variance of  $V$  can be obtained with  $P_{\mathbf{v}\mathbf{v}}$ , the power spectrum of  $\mathbf{v}$ , through

$$\sigma_V^2 = \frac{1}{(2\pi)^3} \int P_{\mathbf{v}\mathbf{v}} \tilde{W}^2 d^3k. \quad (6)$$

Given that radius of the window function  $W(R)$  deployed to measure bulk flow is fairly large, one would expect that the velocity field smoothed at such scale will be well described by linear evolution of the initial condition, then if the initial distribution of velocity is Gaussian, for instance in the case of our simulation, Eq. 5 shall be a good approximation to the PDF of amplitudes of bulk flows. With the model, the most likely amplitude of bulk flow is simply  $V_p = \sqrt{2/3} \sigma_V$ , which

ranges of variance corresponding to different levels can be computed by the integral

$$\int p(V)dV = \text{erf}\left(\sqrt{\frac{3V^2}{2\sigma_V^2}}\right) - \sqrt{\frac{6V^2}{\pi\sigma_V^2}} \exp\left(-\frac{3V^2}{2\sigma_V^2}\right). \quad (7)$$

If  $\mathbf{v}$  is curl free, a velocity potential field can be defined as  $\theta(\mathbf{r}) = -\nabla \cdot \mathbf{v}/(Ha\beta)$  with the scale factor  $a = (1+z)^{-1}$  and the growth rate  $\beta = d\log D/d\log a \approx \Omega_m^{4/7} + (1 + \Omega_m/2)\Omega_\Lambda/70$  (Hamilton 2001), we have

$$\sigma_V^2 = \frac{(Ha\beta)^2}{(2\pi)^3} \int \frac{P_{\theta\theta}}{k^2} \tilde{W}^2 d^3k. \quad (8)$$

The above expression relies on the assumption of negligible rotational velocity. In the linear regime, if the biasing of halo velocity to dark matter velocity is unity, it can be further simplified with the approximation  $P_{\theta\theta} \approx P_{\delta\delta}^{(L)}$  where  $P_{\delta\delta}^{(L)}$  is the linear matter power spectrum.

### 3. BULK FLOW SHOWN IN THE PANGU SIMULATION

#### 3.1. The Pangu simulation and its halo catalogue

The Pangu simulation (PS-I) is a large volume and high resolution simulation, carried out under the scheme of the Computational Cosmology Consortium of China (dubbed C4). PS-I assumes a  $\Lambda$  cold dark matter ( $\Lambda$ CDM) cosmology model with parameters

$$\Omega_m = 0.26, \Omega_b = 0.044, \Omega_\Lambda = 0.74, \\ h = 0.71, \sigma_8 = 0.8, n_s = 1.$$

The simulation contains dark matter only, and uses  $N = 3,072^3$  particles to follow the distribution and evolution of dark matter within a periodic box with  $L = 1000h^{-1}\text{Mpc}$  on a side. Each particle has a mass of  $2.48915 \times 10^9 h^{-1}M_\odot$ . The Plummer-equivalent force softening length is kept constant of  $7h^{-1}\text{kpc}$ .

The PS-I starts from redshift  $z_{\text{init}} = 127$ , initial positions and velocities of particles is generated with Zel'dovich approximation from a glass-like particle set. Input linear power spectrum is computed with the CAMB (Lewis et al. 2000). The simulation is then run with L-GADGET, a memory-optimized version of GADGET2 (Springel 2005). L-GADGET2 is designed to meet requirements of high performance computations, only the tree-particle mesh algorithm is included to calculate the gravitational forces efficiently. Totally 64 snapshots are saved from  $z_{\text{init}} = 127$  to redshift  $z = 0$ . The PS-I is performed on the supercomputer Lenovo Deepcomp7000 at Supercomputing Center of Chinese Academy of Sciences. We use 2048 cores and about 6TB memory at the peak time. The simulation consumes approximately  $6.5 \times 10^5$  CPU hours (about 13 days) in total and consists of 6151 time steps.

Dark matter halos are identified on the fly during the simulation for each snapshot, using the standard friends-of-friends (FOF) algorithm with linking length of 0.2 times the mean particle separation. Each FOF group must contain at least 20 particles, at redshift  $z = 0$  there are  $4.2 \times 10^7$  identified particle groups. There are a little bit of ambiguity in definition of halo mass. Halo mass mostly used in literature is defined as the mass enclosed by a sphere centered on the halo center with certain radius, within which the average density is some factor larger than the critical density. A more convenient definition of halo mass is the total mass of all member particles of the FoF group, which is used in this paper. For simplicity

the average of velocities of member particles are taken as the velocity of the halo.

In this work we used a subset of the group catalogue as our full halo catalogue, which consists of FoF groups of mass larger than  $2.5 \times 10^{11} h^{-1}M_\odot$  only and makes  $1.28 \times 10^7$  entries in the end. Selected FoF groups contains at least 100 member particles, corresponding Poisson fluctuation is greater than 10 which is the threshold used for FoF group identification, limiting discreteness error to 10% level. Another reason is that it is not easy to detect large number of low mass halos in observation. Our prudence test verifies that including FoF groups with number of particle  $< 100$  does not make any significant amendment to final results, even though effects of non-linearity are stronger. More importantly we will show that statistics of bulk flow are not sensitive to the mass of halos in the sample (subsection 3.4).

#### 3.2. The probability distribution function of bulk flow

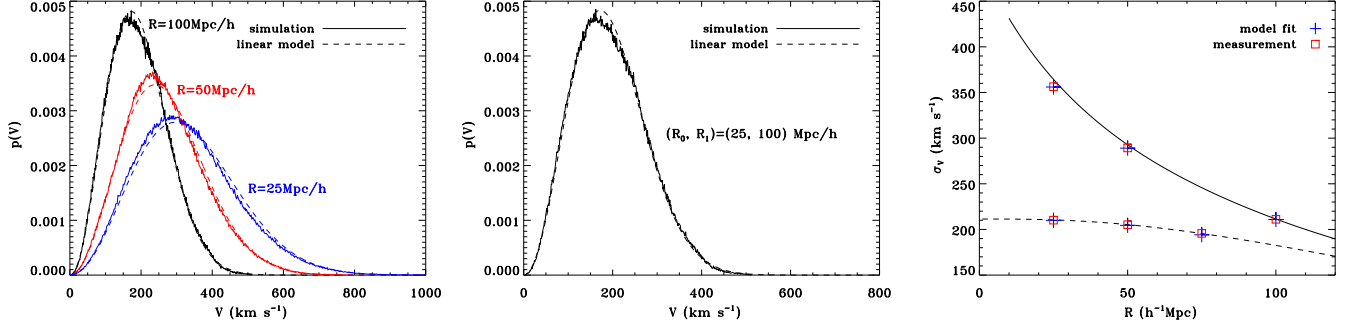
To ensure fair sampling to the simulation, the over-sampling algorithm of Szapudi (1998) is implemented to generate  $\sim 10^6$  cells for  $R$  within  $25 - 100h^{-1}\text{Mpc}$ , bulk flows of halos in random cells are estimated with Eq. 1. In our measurements we adopt mainly the spherical top-hat window, i.e.  $w = 1$  for all halos inside window and  $w = 0$  otherwise, meanwhile shell window function is also deployed for consistency check. Adopting the type of top-hat window function is just to simplify computation, in principle one could try Gaussian or other more sophisticated window functions, but it will not introduce change to the main results. Since our simulation box is limited in a cubic box of side length  $1h^{-1}\text{Gpc}$ , probe of bulk flows in cells of radius  $> 100h^{-1}\text{Mpc}$  would be statistically unreliable.

Our measurements of bulk flows are displayed in Figure 1. For top-hat window,  $\sigma_V$  decreases with cell radius, the possibility of finding extremely large speed of bulk flow becomes smaller for larger volume. For shell window defined by two radius  $R_0 < R_1$ , if  $R_0$  is not very close to  $R_1$ ,  $\sigma_V(R_0, R_1) \sim \sigma_V(R_1)$ .

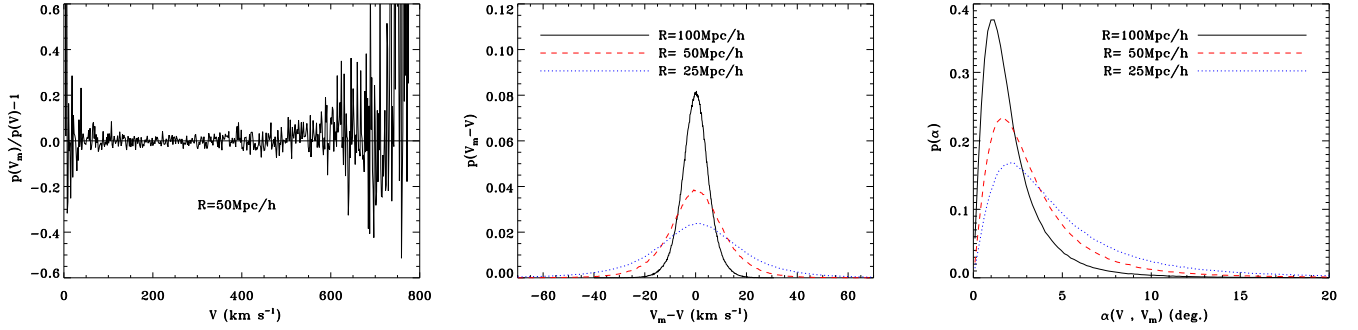
More importantly, PDFs of amplitude of bulk flow in simulation, no matter measured with spherical top-hat window or spherical shell window, are all well described by the linear model. Agreement between simulation and model is better for larger volume as expected, linear theory slightly over-predicts  $\sigma_V$  since nonlinear  $P_{\theta\theta}$  is lower than  $P_{\delta\delta}^{(L)}$  at  $k \sim 0.1h\text{Mpc}^{-1}$  by  $\sim 20\%$  already (Scoccimarro 2004). The comparison clearly lead to the conclusion that to a good precision  $p(V)$  obeys Maxwellian distribution which is completely determined by  $\sigma_V$ , what really matters is not the exact shape of the widow function but rather the corresponding  $\sigma_V$ . This lays out the solid ground for us to put different kinds of measurement together for comparison.

#### 3.3. Bulk flow as mass weighted average of halo velocities

It is known that attenuation to CMB temperature resulted from kSZ effect  $\Delta T_{\text{kSZ}} \propto \mathbf{v} \cdot \hat{\mathbf{l}} \int n_e dl$  in which  $\hat{\mathbf{l}}$  is the unit vector of the line-of-sight and  $n_e$  is the density of free electrons in the galaxy cluster. If the aperture used to measure kSZ effect is sufficiently large, and the number of hot electrons in the cluster can be taken for granted proportional to the mass of host halo  $m_h$ , the total kSZ effect induced temperature fluctuation will be proportional to  $m_h \mathbf{v} \cdot \hat{\mathbf{l}}$ , thus the bulk flow estimated via kSZ effect of galaxy clusters in fact is in principle



**Figure 1.** Left panel:  $p(V)$  measured in spherical top-hat windows with different radius, solid lines are measurements from N-body simulation, from left to right with decreasing heights corresponds to radius of window function  $R = 100, 50, 25 h^{-1} \text{Mpc}$  respectively, dashed lines are prediction of the model of Eq. 5 & 8 with  $P_{\theta\theta} = P_{\delta\delta}^{(L)}$ . Middle panel:  $p(V)$  measured in one spherical shell window defined by two radius  $(R_0, R_1) = (25, 100) h^{-1} \text{Mpc}$ . Right panel: comparison of  $\sigma_V$  predicted by the linear theory (solid line is of spherical top-hat window, dashed line is of spherical shell window with  $R_0 = R, R_1 = 100 h^{-1} \text{Mpc}$ ) with estimation from simulation (blue crosses), and the  $\sigma_V$  which provide the best fitting to PDFs of bulk flow in simulation with Eq. 5 (red squares).



**Figure 2.** Differences between  $\mathbf{V}$  and  $\mathbf{V}_m$  in simulation.  $R$  is the radius of the spherical top-hat window function. Left: the relative difference between  $p(V_m)$  and  $p(V)$ , the large fluctuation at the large  $V$  is due to the almost zero values of PDFs at tail. Middle: PDFs of amplitude difference between  $\mathbf{V}$  and  $\mathbf{V}_m$  in the same cell. Right: distribution functions of the angle between  $\mathbf{V}$  and  $\mathbf{V}_m$  in the same cell,  $\alpha = \cos^{-1}[\mathbf{V} \cdot \mathbf{V}_m / (V V_m)]$ .

the mass weighted average of halo velocities,

$$\mathbf{V}_m = \frac{\sum_i m_i \mathbf{v}_i}{\sum_i m_i}. \quad (9)$$

$\mathbf{V}_m$  is ratio of two Gaussian random variables, the total mass  $M = \sum_i m_i$  and momentum  $\mathbf{P} = \sum_i m_i \mathbf{v}_i$ , with the results of Pham-Gia et al. (2006) it is possible to work out a linear theoretical model for  $p(V_m)$ . Exact calculation needs knowledge of power spectra of matter, momentum and the correlation function between matter and momentum, which is not an easy task. However a quick inspection could give us a rough profile.  $\mathbf{V}_m$  can be expressed as  $(\mathbf{P}/\langle M \rangle)/(\langle M \rangle)$ , i.e. numerator and denominator both are normalized by the mean mass enclosed by the window function, the variance of  $\mathbf{P}/\langle M \rangle$  is dominated by the  $\sigma_V^2 = (2\pi)^{-3} \int P_{\mathbf{v}\mathbf{v}} \tilde{W}^2 d^3k$  if the smoothing scale is sufficiently large (Park & Park 2006), while variance of  $M/\langle M \rangle$  is determined by  $\sigma_\delta^2 = (2\pi)^{-3} \int P_{\delta\delta} \tilde{W}^2 d^3k$ . Since  $\sigma_V^2 \gg \sigma_\delta^2$ ,  $p(M/\langle M \rangle)$  behaves like a Dirac  $\delta$ -function compared with  $p(\mathbf{P}/\langle M \rangle)$ ,  $M$  can then be deemed roughly as a constant and  $p(V_m) \approx \langle M \rangle p(\mathbf{P})$ .  $\mathbf{P}$  is a Gaussian random variable with variance  $\sigma_{\mathbf{P}}^2 \sim \langle M \rangle^2 \sigma_V^2$  in linear regime, therefore  $p(\mathbf{P}) \sim p(V)/\langle M \rangle$  and then  $p(V_m) \sim p(V)$ . Our results of simulation data indeed reveal that there is little difference between  $p(V_m)$  and  $p(V)$  (Fig. 2).

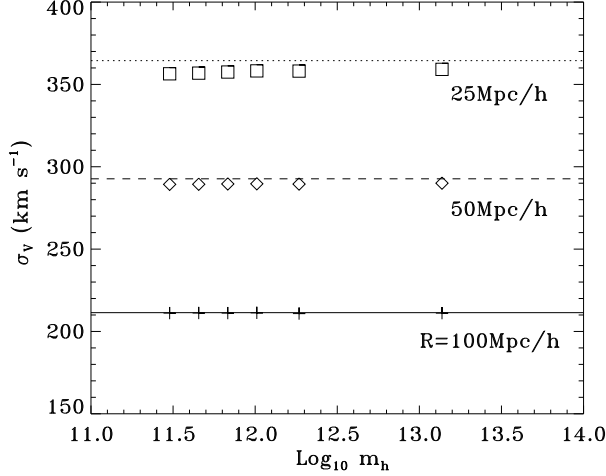
But in an individual cell  $\mathbf{V}_m$  does differ from  $\mathbf{V}$ , both in direction and amplitude. As we can see in Fig. 2,  $p(V_m - V)$  has width of several tens km s<sup>-1</sup> which decreases with larger cell

volume. If the sample volume is small, it could appear that  $V_m$  deviates from  $V$  by  $\sim 100 \text{ km s}^{-1}$  though the possibility is tiny. We also notice that  $V_m - V$  does not show apparent trend with  $V$  or  $V_m$ . Pointing of mass weighted bulk flow does not coincide with  $\mathbf{V}$ . The most likely angle between them is around 1–3 degrees for top-hat window of  $R \in (25, 100) h^{-1} \text{Mpc}$ , and becomes smaller for larger volume. Note that the distribution of the difference angle has a rather long tail, for instance if  $R = 50 h^{-1} \text{Mpc}$  the probability of misalignment greater than  $10^\circ$  is  $\sim 7.3\%$ , yet not trivial.

Hitherto only the kSZ measurements can provide estimation of mass-weighted bulk flow, meanwhile mass weighting does not introduce significant statistical differences, which is actually supported by real observation (Lavaux et al. 2012), thus hereafter we will just concentrate on the unweighted bulk flow.

### 3.4. Halo mass dependence

There is the possibility that bulk flow may depends on the typical mass of halo sample. The full halo catalogue is then divided into six subsamples by halo mass, measured  $\sigma_V$  is plotted in Figure 3 as function of the mean halo mass of the subsample. If the smoothing scale is large enough there is little dependence on mass of sampled halo of bulk flow, but if the window is too small to suppress non-linearity, e.g. the case of  $R = 25 h^{-1} \text{Mpc}$ , subsamples of low mass halos demonstrate slightly lower  $\sigma_V$  than high mass halo subsamples, indicating that low mass halos suffer of more nonlinear action than high



**Figure 3.** Halo mass dependence of  $\sigma_V$ . Symbols are measurement, their x-axis coordinates are the mean halo mass of subsamples. Lines are prediction of linear theory, below which are labelled with corresponding cell radius.

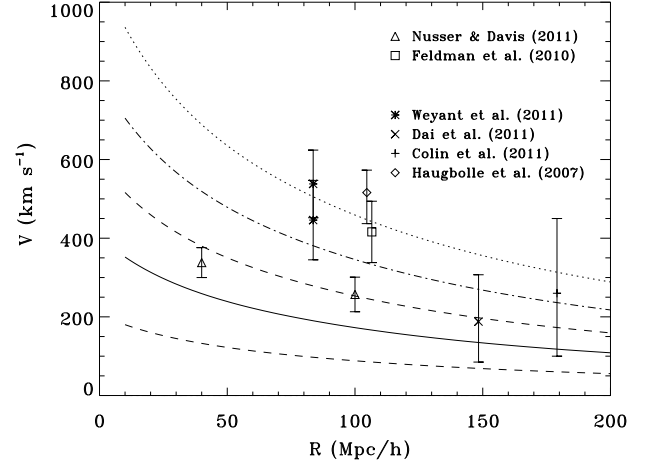
mass halos.

However for an individual volume, diversity in bulk flows given in different halo mass bins is very likely to be apparent when  $V \lesssim \sigma_V$ , although it declines with increasing  $V$  and turns out insignificant for  $V \gtrsim 2\sigma_V$ . Our experiments disclose that differences between bulk flows of the lightest halo subsample ( $\bar{m}_h = 3.015 \times 10^{11} h^{-1} M_\odot$ ) and the mostly massive halo subsample ( $\bar{m}_h = 1.374 \times 10^{13} h^{-1} M_\odot$ ) in numerical figures is actually extremely similar to Figure 2. Therefore if sampled volume in observation has low depth, for instance the effective  $R < 50 h^{-1} \text{Mpc}$ , and measured bulk flow is not of extremely high amplitude, attempt to find rigorous convergence among bulk flows estimated from different galaxy or galaxy cluster samples could be probably fruitless.

### 3.5. Consistency between observation and model

Systematical biases in dark flow, the bulk flow measured at high redshift, are not fully understood and precisely controlled, so we refrain ourselves from discussing high redshift case. Most of the local (or nearby) bulk flow measurements has redshift less than  $\sim 0.06$ , resulting redshift evolution of  $\sigma_V$  with respect to  $z = 0$  is of magnitude of a few percents at most, which can be comfortably ignored. Window functions in different works are not the same at all, but the excellent performance of the linear model provides an unified scheme. Since PDFs of bulk flow is solely determined by  $\sigma_V$ , independent of the type of the window, the radius of a top-hat window which gives the same linear  $\sigma_V$  as the window function used in observation can acts as the effective scale corresponding to a particular sample.

To check if an observed bulk flow is consistent with  $\Lambda$ CDM model, we need to figure out the variance ranges of  $V$ . The most likely amplitude of  $V$  is  $V_p = \sqrt{2/3}\sigma_V$ , derived via  $dp(V)/dV = 0$ , and the variance range at different levels are computed through Eq. 7. The Maxwellian distribution is highly skewed, the one side probability  $\int_0^{V_p} p(V)dV = 42.8\%$ , but confidence probabilities at levels of  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  are  $\epsilon_{1,2,3} = 68.3\%, 95.5\%, 99.7\%$  respectively, so variance ranges at  $2\sigma$  and  $3\sigma$  significance are defined by  $(0, V_{2,3})$  with  $\int_0^{V_{2,3}} p(V)dV = \epsilon_{2,3}$ , while the  $1\sigma$  variance is within  $(V_1^-, V_1^+)$



**Figure 4.** Measured local bulk flows against prediction of linear model. Symbols are some recently estimated amplitudes of local bulk flows from observational data of galaxies (Feldman et al. 2010; Nusser & Davis 2011) and supernovae (Haugbølle et al. 2007; Colin et al. 2011; Dai et al. 2011; Weyant et al. 2011). Solid line is the model predicted most likely bulk speed, dashed lines indicate variance at  $1\sigma$  level (68.3%), dot-dashed line and dotted line are at levels of  $2\sigma$  (95.5%) and  $3\sigma$  (99.7%) respectively.

given  $\int_{V_1^-}^{V_p} p(V)dV = \int_{V_p}^{V_1^+} p(V)dV = \epsilon_1/2$ .

Several recent measurements from observational data are over-plotted upon model prediction in Fig. 4. One has to keep in mind that our calculation of effective scales has assumed that all other factors affecting estimation have been perfectly corrected, such as non-linearity, selection function and sky coverage incompleteness. Construction details of these samples are often too sketchy to render appropriate weights for our calculation. However the imperfection actually reduces their effective volumes so that the true effective scales will be smaller, i.e. the data points will shift leftward along horizontal axis in Fig. 4. So we shall deem Fig. 4 as the mostly conservative judgment of the consistency between observation and theory. Nonetheless, from Figure 4, it appears that observation of our local Universe does not rule out the  $\Lambda$ CDM model. Those results of Haugbølle et al. (2007), Feldman et al. (2010) and Weyant et al. (2011) that often quoted as supporting evidence disfavoring standard  $\Lambda$ CDM model are around  $3\sigma$  level, but the significance will be smaller if error bars are taken into account. In addition, considering that many other measurements (including the not officially published report of Wang 2007) are in fact consistent with  $\Lambda$ CDM model, we prefer to choose conservative standpoint on the issue.

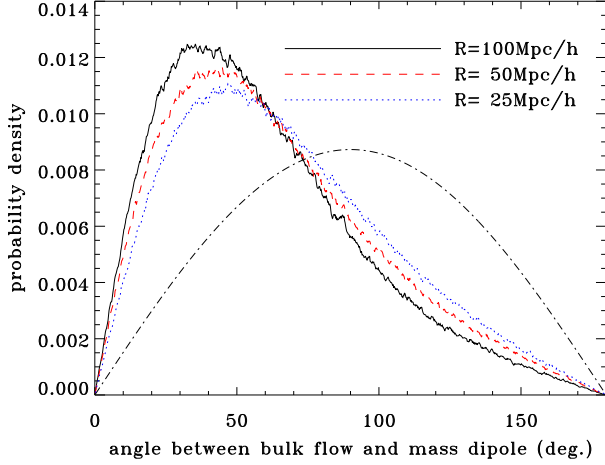
### 3.6. Bulk flow and mass distribution in the cell

It is interesting to investigate the relation between bulk flow and the mass distribution in the sample volume, one might wonder whether one could infer bulk flow from mass distribution if peculiar velocity data is absent, since in linear regime Fourier modes of velocity field can be derived from modes of the density field. However from Eq. 2 it is clear that bulk flow is determined by those modes of wavelengths larger than the characteristic scale of the window function, if the Fourier transformation of the density field is restricted to the same volume in which bulk flow is measured, those modes of long wavelength accounted for bulk flow are missing.

Our measurements confirm the expectation. The first quantity we checked is the mass monopole, the total mass ( $\sum_i m_i$ )

**Table 1**  
Correlation coefficients between amplitude of bulk flow and mass distribution

Cell radius	Mass monopole		Mass dipole		
$R$ ( $h^{-1}$ Mpc)	$\sum_i m_i$	$N$	$\frac{\sum m_i \mathbf{r}_i / r_i}{\sum m_i}$	$\langle \frac{\mathbf{r}_i}{r_i} \rangle$	
25	0.024	-0.0004	...	0.139	0.134
50	0.004	-0.0188	...	0.188	0.187
100	0.014	-0.0031	...	0.245	0.249



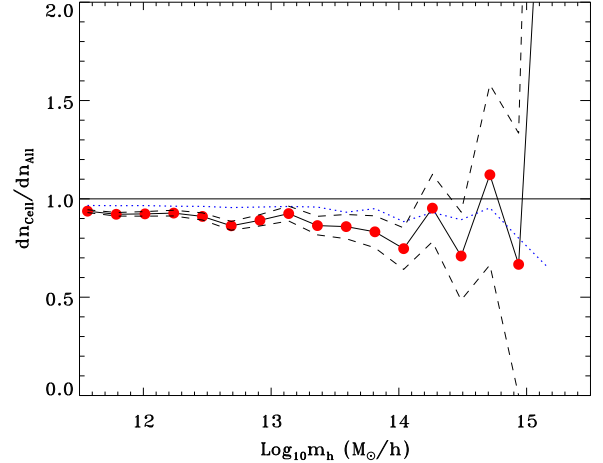
**Figure 5.** Probability distribution function of the angle between bulk flow and mass dipole. Dash-dot line is the expectation of null correlation.

or the total number ( $N$ ) of halos in the volume, which is equivalent to the density fluctuation smoothed by the window function. Correlation coefficients are computed to denote the correlation strength between amplitudes of bulk flow and mass monopole (Table 1). Apparently bulk flow is not correlated with the total mass and the total number of halos in the volume at all.

Two kinds of dipoles of halo distribution are measured,  $\langle \mathbf{r}_i / r_i \rangle$  and the mass weighted one ( $\sum m_i \mathbf{r}_i / r_i / \sum m_i$ ). Our results show that including halo mass or not makes little difference. As can be seen in Table 1 and Figure 5, mass dipole correlates with bulk flow very weakly both in amplitude and direction. The correlation becomes slightly tighter as sample volume increases, the peak of the distribution function of the misalignment angle between mass dipole and bulk velocity is at  $\sim 45^\circ$  for  $R = 25 h^{-1}$  Mpc while shifts to  $\sim 32^\circ$  for  $R = 100 h^{-1}$  Mpc (Figure 5).

#### 4. THE FASTEST BULK FLOW

There are some works claim detection of unusually large bulk flow (e.g. Feldman et al. 2010; Weyant et al. 2011), it is interesting to check properties of these special parts in  $\Lambda$ CDM universe. The largest amplitudes of bulk motion measured in our simulation for  $R = 25, 50, 100 h^{-1}$  Mpc are 1070, 778, 514  $\text{km s}^{-1}$  respectively, the one for  $R = 100 h^{-1}$  Mpc is already very close to those observational results. The possibility of residing in the cell having the fastest bulk motion is defined by the ratio of the cell volume to the total volume of simulation,  $\approx 0.42\%$  for  $R = 100 h^{-1}$  Mpc. However, given the diameter of the cell as large as  $200 h^{-1}$  Mpc against



**Figure 6.** Difference in halo abundance between the fastest cell with  $R = 100 h^{-1}$  Mpc and the full halo catalogue,  $dn_{\text{Cell}}/dn_{\text{All}}$ ,  $dn_{\text{Cell}}$  is the mass function of halos in the fastest cell,  $dn_{\text{All}}$  is of the full halo catalogue. Red solid circles connected by solid line is of the fast cell, which is enclosed by dashed lines marking its Poisson variance. The blue dotted line is the average of the top-ten fastest cells. The horizontal solid line is of  $dn_{\text{Cell}}/dn_{\text{All}} = 1$ .

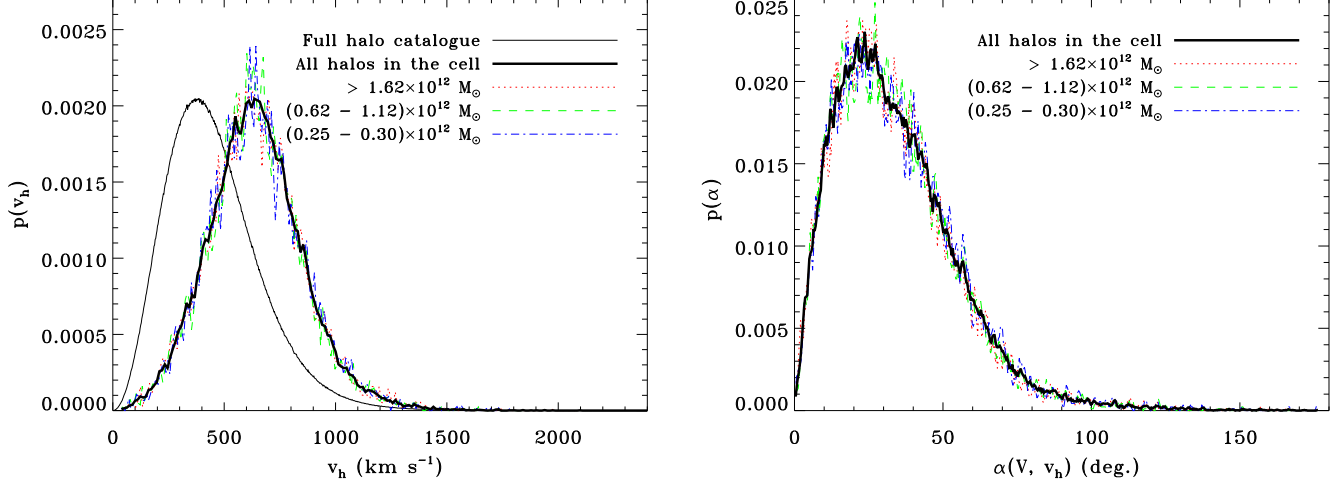
the simulation box length  $1 h^{-1}$  Gpc, a huge volume moving at speed more than  $\sim 500 \text{ km s}^{-1}$  will yield observable features too prominent to be missed.

In this report we will choose the  $R = 100 h^{-1}$  Mpc case as example to study peculiarity of the cell showing fastest bulk motion. The first physical quantity checked is the halo mass function in the cell, for comparison halo mass functions of the top-ten fastest cells (centers separated by at least  $100 h^{-1}$  Mpc) are also measured. Shown in Figure 6 are measured cell halo mass functions divided by the halo abundance in the full catalogue. Among the ten cell mass functions, most (more than 7 of 10) are smaller than the halo mass function of the full catalogue for  $m_h < \sim 10^{14} h^{-1} M_\odot$ . In high mass regime due to the very small number of high mass halos we can not withdraw any reliable conclusions though the mass function of the fastest cells demonstrates a high tail. So far we would only cautiously conclude that in high bulk flow regions there is the tendency of finding less number of small mass halos.

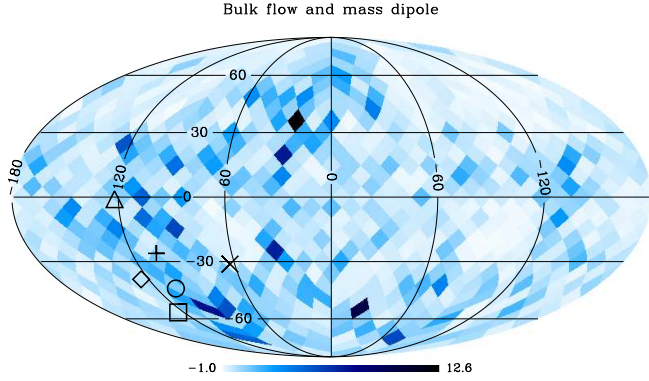
Distribution of all halo velocities in the cell is illustrated in Figure 7, also shown are velocity distribution functions of these halos in mass bins  $(0.25 - 0.3) \times 10^{12} h^{-1} M_\odot$ ,  $(0.62 - 1.12) \times 10^{12} h^{-1} M_\odot$  and  $> 1.62 \times 10^{12} h^{-1} M_\odot$ . Mass binned halos do not exhibit any significant differences in aspect of velocity distribution, which eases the worry of possible bias in mass selected halo samples. Distribution function of the angle between halo velocity and bulk flow is very skewed toward small misalignment, the peak is around  $20 \sim 30^\circ$  but not the  $0^\circ$ , about 90% halos are moving in direction within  $60^\circ$  to the bulk flow. It appears that halos in the cell with largest bulk velocity are more likely to have higher speed, the peak of velocity amplitude distribution of halos in the cell is at around  $600 \text{ km s}^{-1}$  while that of the full halo catalogue is at  $\sim 380 \text{ km s}^{-1}$  (Figure 7).

It has been examined that bulk flow basically is weakly correlated with the internal mass dipole on average. But for the cell with the largest bulk flow, intuitively one would conjecture there should be certain very massive clumps neighboring to the cell, their gravitational action may play a dominant roll in causing such extreme bulk flow of nearby halos. As an at-





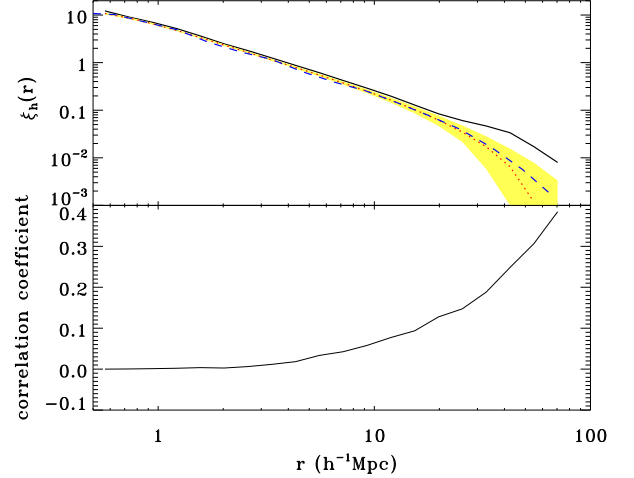
**Figure 7.** Velocity distribution of halos in the cell with the largest bulk flow ( $R = 100h^{-1}\text{Mpc}$ ). Left panel is the distribution of amplitudes of halo velocities, thick solid line is of halos in the cell, thin solid line is of the full halo catalogue, other color lines are of halos in three different mass bins as labelled in figure legend. Right panel presents distributions of the angle between bulk flow  $\mathbf{V}$  and halo velocity  $\mathbf{v}_h$ ,  $\alpha = \cos^{-1}[\mathbf{V} \cdot \mathbf{v}_h / (Vv_h)]$ .



**Figure 8.** Mollweide projection of directions of the largest bulk flow and mass dipoles. Mass dipoles are computed from the mass field represented by halos,  $\sum m_i \mathbf{r}_i / r_i$ . The plus symbol marks the direction of the largest bulk flow ( $514 \text{ km s}^{-1}$ , cell radius  $R = 100h^{-1}\text{Mpc}$ ), the cross symbol is the direction of dipole of halos in the cell, the angle between them is  $37.4^\circ$ . Triangle, square, diamond, and circle symbols are directions of mass dipoles within shells (100, 150), (150, 200), (200, 250), (250, 300)  $h^{-1}\text{Mpc}$ , which deviate from the bulk flow by about  $30^\circ$ ,  $33^\circ$ ,  $21^\circ$ ,  $17^\circ$  respectively. The color map is the surface mass density contrast of halos in the cell, plotted in the Healpix scheme (dark color for high density contrast).

tempt to justify the paradigm, mass dipoles in shells within  $R = 100 - 300h^{-1}\text{Mpc}$  to center of the  $100h^{-1}\text{Mpc}$  cell with largest bulk flow are calculated in four layers with the help of the Healpix package (Górski et al. 2005), if there is unusual distribution of matter in a layer, mass dipole of the layer will be the efficient indicator. Projected directions of bulk flow and mass dipoles are displayed in Figure 8. When shell moves outward mass dipole pointing walks fairly randomly around the bulk flow, the misalignment angle varies between  $\sim 20 - 40^\circ$  which is analogous to the typical value in Figure 5. It seems that dipoles of local environmental mass are only aligned crudely with the bulk flow. The correlation is not negligible, however since these misalignment angles are not small, we have no strong support from the simulation to attribute the extremely large bulk flow in such a huge volume mainly to inhomogeneous environment.

Anomaly in clustering of halos in the special cell is detected



**Figure 9.** Correlation between bulk velocity and clustering of halos in the cell, cell radius is  $R = 100h^{-1}\text{Mpc}$ . Top panel: solid line is the two-point correlation function  $\xi_h$  of halos in the cell with the largest bulk velocity, dotted line in red is the average over 1000  $\xi_h$ s measured in randomly selected cells while the yellow shadow marks the corresponding  $1\sigma$  variance, dashed line in blue is the result of all halos in our halo catalogue. Bottom: correlation coefficients between amplitude of bulk velocity and  $\xi_h(r)$  as function of  $r$ , summarized from measurements of 1000 randomly selected cells.

(Figure 9), halo two-point correlation functions of 1000 randomly located cells of radius  $100h^{-1}\text{Mpc}$  and the full halo catalogue are also calculated.  $\xi_h$  averaged over the 1000 measurements agrees with  $\xi_h$  of the full halo catalogue at scales  $r \lesssim 30h^{-1}\text{Mpc}$ , then drops down more quickly to zero at larger scales due to the integral constraint resulted from finite volume of cell (Landy & Szalay 1993). As we are interested in  $\xi_h$  in a finite volume, we did not bother ourselves to apply relevant correction.  $\xi_h$  of the cell with largest bulk velocity is higher than the average of random cells by around  $2\sigma$  at scales  $r \lesssim 20h^{-1}\text{Mpc}$ , at larger scales the excess of clustering power rises to level of  $\sim 2 - 3\sigma$ . In order to assess the statistical significance of the event, correlation coefficients of bulk flow amplitude and  $\xi_h(r)$  are computed from the 1000

random cells (bottom panel in Figure 9). Little correlation is detected at scales  $r \lesssim 20h^{-1}\text{Mpc}$ , then at larger scales the correlation becomes much stronger, indicating that in cells with high bulk velocity it is truly more possible to find power excess in halo clustering at large scales, which appears to be in line with the findings in Macaulay et al. (2011). We further check the halo two-point correlation functions in the top-ten fastest cell, and we find that 7 of the 10 demonstrate power excess at similar scales. Note that the power excess at large scales can not be ascribed to integral constraint, for the leading term of integral constraint in this regime is negative and will bend  $\xi_h$  downward (Landy & Szalay 1993).

## 5. SUMMARY AND DISCUSSION

Through analysis of the Pangu simulation, it is confirmed that bulk flow of halos follows Maxwellian distribution which is completely determined by a single parameter, the bulk velocity dispersion. We find that the dispersion measured in simulation agrees with the prediction of linear perturbation theory of structure formation very well, non-linearity only becomes important when the sampling volume is very small. We find that in most cases mass weighted bulk flow has some minor statistical differences to the unweighted one, but the will not affect the overall statistics significantly. It is also revealed that statistically bulk flow has little systematical dependence on the mass of halos used for estimation. Based on the results, we propose a unified scheme to compare results from observational samples with theories. In the proposal, the scale at which bulk flow in a particular space of the Universe is measured is chosen to be the effective scale  $R$  which is the radius of a spherical top-hat window function  $W_h$  that yields the same bulk velocity dispersion  $\sigma_V$  as the practical window function  $W_O$  for the observational sample does in linear theory.  $W_O$  is not only determined by the sample geometry but also contains weights emerged from selection function, incompleteness and etc., being analogous to the window function used in estimation of power spectrum.

Variance ranges of bulk flow are clarified as well on the basis of Maxwellian distribution in the work, we make a rough comparison of some recently measurements with the  $\Lambda\text{CDM}$  model adopted in Pangu simulation, we find that part results do deviate from the model by about  $3\sigma$  but the tension between observation and model is not so strong as original works claimed. Estimated effective scales for observation results in Figure 4 are in fact the upper limits, the true effective scales could be even smaller since we have assumed that those samples are of full sky coverage and their selection functions have been corrected for during estimation. Furthermore, observed bulk velocity consists of residuals from thermal motion of galaxies in their host halo, which is not included in calculation of the velocity dispersion so far. More accurate modeling could be developed by assuming velocities of galaxies relative to their halos obey certain simple distribution, but it requires explicit knowledge of occupation details of galaxies in host halo, which in itself is already a challenging problem. A better way would be to deduct the random motion component in the estimation procedure, such as the treatment in Wang (2007).

Correlation between bulk flow and dipole of internal mass is very weak, but is stronger for larger volume. If one happens to be living in a volume of radius greater than  $100h^{-1}\text{Mpc}$  with large bulk velocity, their observed mass dipole in the volume will have considerable chance of being unusually strong. Typical misalignment angle between bulk flow and mass dipole is

mostly likely around  $\sim 30-50^\circ$ . This might introduce non-negligible systematical bias to cosmological probes involving local mass distribution, such as the late-time integrated Sachs-Wolfe effect (Rees & Sciama 1968).

In our simulation there do exist volume of scale extending to  $200h^{-1}\text{Mpc}$  in diameter moving with extreme large bulk velocity more than  $500\text{kms}^{-1}$ . Most halos inside the volume are moving in alignment with the bulk flow within  $60^\circ$ , and the flow shows no dependence on halo mass. Such group motion of numerous halos will generate prominent kSZ signals, it is a rare event, but given its high speed and huge scale ( $200h^{-1}\text{Mpc}$  versus  $1h^{-1}\text{Gpc}$ ), probability of detection is actually not very small, which of course also depends on the inclination between bulk flow and line-of-sight. Another possible observation effect is that galaxies in the volume could be dimmed or brightened on average by the extreme bulk flow than galaxies in other places, which is the starting point of the effort tried by Nusser et al. (2011) and Abate & Feldman (2012).

Dipoles of mass outside the largest bulk flow region as environment are not tightly aligned with the bulk velocity, but deviate from it by around  $\sim 20-40^\circ$ . Interestingly we identified that halo clustering of the particular volume is strengthened apparently at scales  $r \gtrsim 20h^{-1}\text{Mpc}$ , simulation results point out that such enhancement is not completely accidental, at large scales two-point correlation function of halos in a finite volume is indeed mildly correlated with bulk velocity. Bulk velocity is dominated by Fourier modes of velocity at scales larger than the characteristic scale of the sample while in linear theory  $\mathbf{v}(\mathbf{k}) \propto i\mathbf{k}\delta(\mathbf{k})/k^2$ , unusually large bulk velocity seems to imply that there should be extraordinary super large mode of density fluctuation topping up in the region. However bulk flow is hardly correlated with the total number or mass of enclosed halos inside the sample volume, and as we checked the total number or mass of halos in the cell with largest bulk velocity is less than the mean value but still within  $2\sigma$  variance range. Moreover, in linear regime Fourier modes of density fluctuation are independent,  $\xi$  of halos inside the volume is controlled actually by modes of scale less than the characteristic scale of the sample.

Aside from the theoretical puzzle, one question is whether the power excess of halo clustering in a region at scales  $\gtrsim 20h^{-1}\text{Mpc}$  can be used as indicator of candidate space of extremely large bulk flow, the advantage of using two-point correlation function is that clustering does not rely on direction of line-of-sight, the complication resulted from redshift distortion in principle can be overcome by the ratio of  $\xi_h(r \sim 20-50h^{-1}\text{Mpc})$  to  $\xi_h$  at small scales e.g.  $\sim 10h^{-1}\text{Mpc}$  where  $\xi_h$  is not correlated with bulk flow. A more serious concern is that if we are unlucky (or lucky) in a special region as large as our current largest galaxy survey with extremely large bulk flow, the measured clustering strength at and beyond scale of baryonic acoustic oscillation would be significantly leveled up, could it be the case of power excess at very large scales in the Baryon Oscillation Spectroscopic Survey (BOSS) elaborated by Ross et al. (2012)? To answer all these queries one surely needs multiple realizations of simulation of volume much bigger than our Pangu simulation, for the moment in this paper we have to leave these questions open.

## ACKNOWLEDGMENT

This work is partly supported by the NSFC through grants of Nos. 10873027, 10873035, 11073055, and 11133003.



YPJ, WPL, XHY and PJZ are members of the Innovation group funded by NSFC (No. 11121062). JP and XK acknowledge the One-Hundred-Talent fellowships of CAS. We appreciate stimulating discussion with Jiasheng Huang, Cheng Li, Guoliang Li and Lifan Wang.

The Pangu simulation was carried out in the Supercomputing center of CNIC, CAS, under the collaboration scheme of the Computational Cosmology Consortium of China (C4), participating institutions are NAOC, PMO, SHAO and CNIC.

## REFERENCES

- Abate, A., & Feldman, H. A. 2012, *MNRAS*, 419, 3482
- Afshordi, N., Geshnizjani, G., & Khoury, J. 2009, *JCAP*, 8, 30
- Amanullah, R., et al. 2010, *ApJ*, 716, 712
- Bahcall, N. A., Cen, R., & Gramann, M. 1994, *ApJ*, 430, L13
- Colin, J., Mohayaee, R., Sarkar, S., & Shafieloo, A. 2011, *MNRAS*, 414, 264
- Courtois, H. M., Tully, R. B., Makarov, D. I., Mitronova, S., Koribalski, B., Karachentsev, I. D., & Fisher, J. R. 2011, *MNRAS*, 414, 2005
- Dai, D.-C., Kinney, W. H., & Stojkovic, D. 2011, *JCAP*, 4, 15
- Feldman, H. A., Watkins, R., & Hudson, M. J. 2010, *MNRAS*, 407, 2328
- Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, *ApJ*, 622, 759
- Hamilton, A. J. S. 2001, *MNRAS*, 322, 419
- Haugbølle, T., Hannestad, S., Thomsen, B., Fynbo, J., Sollerman, J., & Jha, S. 2007, *ApJ*, 661, 650
- Jarosik, N., et al. 2011, *ApJS*, 192, 14
- Jha, S., Riess, A. G., & Kirshner, R. P. 2007, *ApJ*, 659, 122
- Kashlinsky, A., Atrio-Barandela, F., & Ebeling, H. 2011, *ApJ*, 732, 1
- Kashlinsky, A., Atrio-Barandela, F., Ebeling, H., Edge, A., & Kocevski, D. 2010, *ApJ*, 712, L81
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
- Lavaux, G., Afshordi, N., & Hudson, M. J. 2012, arXiv:1207.1721
- Lewis, A., Challinor, A., & Lasenby, A. 2000, *ApJ*, 538, 473
- Macaulay, E., Feldman, H., Ferreira, P. G., Hudson, M. J., & Watkins, R. 2011, *MNRAS*, 414, 621
- Mak, D. S. Y., Pierpaoli, E., & Osborne, S. J. 2011, *ApJ*, 736, 116
- Mersini-Houghton, L., & Holman, R. 2009, *JCAP*, 2, 6
- Moscardini, L., Branchini, E., Brunozzi, P. T., Borgani, S., Plionis, M., & Coles, P. 1996, *MNRAS*, 282, 384
- Nusser, A., Branchini, E., & Davis, M. 2011, *ApJ*, 735, 77
- Nusser, A., & Davis, M. 2011, *ApJ*, 736, 93
- Osborne, S. J., Mak, D. S. Y., Church, S. E., & Pierpaoli, E. 2011, *ApJ*, 737, 98
- Park, C.-G., & Park, C. 2006, *ApJ*, 637, 1
- Pham-Gia, T., Turkkan, N., & Marchand, E. 2006, *Communications in Statistics - Theory and Methods*, 35, 1569
- Rees, M. J., & Sciama, D. W. 1968, *Nature*, 217, 511
- Ross, A. J., et al. 2012, arXiv:1203.6499
- Sandage, A., Reindl, B., & Tammann, G. A. 2010, *ApJ*, 714, 1441
- Scoccimarro, R. 2004, *Phys. Rev. D*, 70, 083007
- Song, Y.-S., Sabiu, C. G., Kayo, I., & Nichol, R. C. 2011, *JCAP*, 5, 20
- Springel, V. 2005, *MNRAS*, 364, 1105
- Springob, C. M., Masters, K. L., Haynes, M. P., Giovanelli, R., & Marinoni, C. 2007, *ApJS*, 172, 599
- Strauss, M. A., & Willick, J. A. 1995, *Phys. Rep.*, 261, 271
- Sunyaev, R. A., & Zeldovich, I. B. 1980, *MNRAS*, 190, 413
- Szapudi, I. 1998, *ApJ*, 497, 16
- Wang, L. 2007, arXiv:0705.0363
- Watkins, R., Feldman, H. A., & Hudson, M. J. 2009, *MNRAS*, 392, 743
- Weyant, A., Wood-Vasey, M., Wasserman, L., & Freeman, P. 2011, *ApJ*, 732, 65
- Wyman, M., & Khoury, J. 2010, *Phys. Rev. D*, 82, 044032